

PATENT ABSTRACTS OF JAPAN

(11)Publication number : **62-274448**

(43)Date of publication of application : **28.11.1987**

(51)Int.CI.

G06F 13/10
G06F 9/46

(21)Application number : **61-117490**

(71)Applicant : **HITACHI LTD**

(22)Date of filing : **23.05.1986**

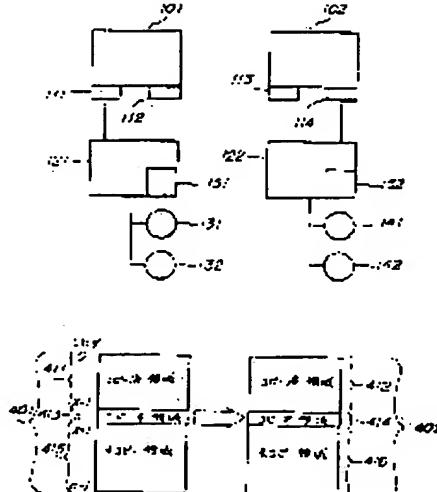
(72)Inventor : **TAKAMATSU HISASHI
KURANO AKIRA
TAKEUCHI HISAHARU
NOZAWA MASASHI
KODAIRA MITSUHIKO**

(54) CONTROL SYSTEM FOR INPUT/OUTPUT DEVICE

(57)Abstract:

PURPOSE: To share a file without reducing the system throughput by making it possible for each OS to change and test control state information as a part of the input/output operation and detecting the change of control state information without testing control state information by the OS.

CONSTITUTION: During the copying operation, the whole of disc devices 131 and 142 is not closed but only current copying areas 413 and 414 of respective disc devices and an uncopied area 416 of a disc device 142 are closed. Consequently, normal input/output operations other than the copying operation are executed by dual operation while the copy operation to copied areas 411 and 412 of both disc devices is performed, and the identify of data contents is kept. Execution of normal input/output operations other than the copying operation to copying areas 413 and 414 is suspended until the copying operation to their cylinders is terminated. Normal input/output operations other than the copying operation to the uncopied area 415 are executed by single operation of only the disc device 131.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

THIS PAGE BLANK (USPTO)

⑨ 日本国特許庁 (JP)

⑩ 特許出願公開

⑪ 公開特許公報 (A)

昭62-274448

⑫ Int.Cl.⁴G 06 F 13/10
9/46

識別記号

330
340

厅内整理番号

C-7165-5B
F-8320-5B

⑬ 公開 昭和62年(1987)11月28日

審査請求 未請求 発明の数 1 (全8頁)

⑭ 発明の名称 入出力装置の制御方式

⑮ 特願 昭61-117490

⑯ 出願 昭61(1986)5月23日

⑰ 発明者 高松 久司 小田原市国府津2880番地 株式会社日立製作所小田原工場内

⑰ 発明者 倉野 昭 小田原市国府津2880番地 株式会社日立製作所小田原工場内

⑰ 発明者 竹内 久治 小田原市国府津2880番地 株式会社日立製作所小田原工場内

⑰ 発明者 野沢 正史 小田原市国府津2880番地 株式会社日立製作所小田原工場内

⑰ 発明者 小平 光彦 秦野市堀山下1番地 株式会社日立製作所神奈川工場内

⑯ 出願人 株式会社日立製作所 東京都千代田区神田駿河台4丁目6番地

⑰ 代理人 弁理士 武頭次郎 外1名

明細書

1. 発明の名称

入出力装置の制御方式

2. 特許請求の範囲

1. 複数のオペレーティング・システムにより構成される情報処理システムにおいて、複数のオペレーティング・システムから共用される入出力装置及び入出力制御装置と、前記入出力装置へのデータの入出力を含む入出力装置の管理状態を管理状態情報として記憶するために、該入出力装置または入出力制御装置内に設けた記憶手段と、記憶された管理状態情報を、入出力動作の一部として前記入出力装置に対する通常の入出力動作に先立つてテストする、前記複数のオペレーティング・システムの各々に設けられたテスト手段とを備えることにより、前記複数のオペレーティング・システムの夫々が、前記入出力装置の管理状態に応じた処理を行い得るようにしたことを特徴とする入出力装置の制御方式。

2. 前記情報処理システムにおいて、入出力装置の管理状態の変更を、接続されているオペレーティング・システムに通知する手段を前記入出力装置または入出力制御装置に備えることを特徴とする前記特許請求の範囲第1項記載の入出力装置の制御方式。

3. 発明の詳細な説明

〔産業上の利用分野〕

本発明は、情報処理システムに用いられる入出力装置を複数のオペレーティング・システムが共用する場合における入出力装置の制御方式に関する。

〔従来の技術〕

情報処理システムは、システムの信頼性や処理能力を向上するため、複数のオペレーティング・システム(以下単にOSという)を用いて構成される場合がある。このとき、ある1個のOSが処理したデータを他のOSで処理する必要性などから、入出力装置を複数のOSで共同して使用することになる。

特開昭62-274448 (2)

このような場合、入出力装置の管理状態がある1個のOSにより変更されても、その変更が直ちに全ての他のOSに伝達されないため、夫々のOSにおける中央処理装置内に記憶されている入出力装置の管理状態情報が一致しなくなる。このため、入出力装置の管理状態変更後に、変更前の管理状態情報を持つ別のOSが、この管理状態を変更された入出力装置に対して入出力命令を実行すると、適切な入出力動作を実行することができないという問題点があつた。

更にまた、共用される入出力装置のある機能に対して管理状態を変更するため、あるOSがその入出力装置に対して初期設定などの複雑な準備を行つている間、すなわち、この準備が完了する迄の間、準備を行つているOS以外のOSからのこの機能の利用ができないようにしておく必要がある。

前述の問題点を回避し、前述の必要性を満たすため、入出力装置の管理状態の変更を行う場合、チャネル間結合装置等を設けて、OS相互間で連

号公報に述べられている。該公報に開示された従来技術は、入出力制御装置にメモリを設置することにより、二重化された入出力装置としてのディスク装置の内容が同一でなくなつたとき、高速に両ディスク装置の内容を同一化するものである。しかし、この従来技術では、複数のOSが共同して入出力装置を使用する場合については、何ら配慮されていない。

〔発明が解決しようとする問題点〕

前述したように、従来技術においては、いずれの方法によつても、複数のOSを用いる情報処理システムにおいて、OSからの入出力動作が適切に実行できない場合が生じたり、また、入出力動作に時間がかかる等の欠点があつた。

本発明の目的は、複数のOSが相互に直接連絡を取り合うことなく、複数のOSが共用する入出力装置の管理状態を、システム・スループットを低下させることなく検知できるようにした入出力装置の制御方式を得ることにある。

〔問題点を解決するための手段〕

格を行うようにする方法がある。しかし、この方法では、連絡に要する時間による遅延が生じ、各OS内の入出力装置の管理状態情報の変更に時間がかかるという欠点を生じる。また、連絡すべきOSがシステム・ダウンしている場合には、適切な連絡が行えないため、特別な配慮が必要となる。

前述したOS相互間で連絡をする代りに、複数のOSで共用されるディスク装置などの外部記憶装置に管理状態情報を記憶しておき、入出力装置への入出力の際に、必ずこの外部記憶装置に記憶された管理状態情報をテストし、該情報に変更がなかつた場合だけ、入出力動作の実行を行うようとする方法もある。しかし、この方法は、入出力装置に対する通常の入出力動作の都度、外部記憶装置から管理状態情報を読出してテストするための時間が加わり実用的でない。

そこで、入出力装置の管理状態情報を入出力装置または入出力制御装置に記憶する方法が考えられる。この種の従来技術が、特開昭59-180765

本発明によれば、前記目的は、複数のOSから共用される入出力装置または入出力制御装置内に、複数のOSから共用される入出力装置の管理状態を管理状態情報として格納記憶する記憶手段を設け、各OSは、入出力動作の一部としてこの管理状態情報の変更およびテストができるようにし、さらに、変更された管理状態情報を入出力装置が接続されているOSに連絡する手段を入出力装置に設け、OSが入出力装置または入出力制御装置内の管理状態情報をテストしなくとも管理状態情報の変更を検知できるようにすることにより達成される。

〔作用〕

本発明における、入出力装置または入出力制御装置に設けられた記憶手段には、対応する入出力装置に関する管理状態情報が記憶されている。各OSは、入出力動作を実行する都度、この記憶手段をテストすることにより、その入出力装置の管理状態の変更を検知し、自己が有する管理状態情報を更新することができるので、他のOSにより

特開昭62-274448 (3)

入出力装置の管理状態が変更されている場合も、管理状態の変更に応じた処理を実行することができる。

〔実施例〕

以下、本発明による入出力装置の制御方式について、入出力装置の具体例としてディスク装置をあげ、図示の実施例により詳細に説明する。

第1図は、複数のOSがディスク装置を共用する場合の本発明の一実施例を示すブロック図、第2図は、ディスク制御装置内に設けられた管理状態情報を格納するメモリのうち、1個のディスク装置に対応するメモリ部分のメモリ・マップの一例を示す図、第3図は、コマンド・チエインを示す図、第4図は、二重化されたディスク装置の内容が不一致となつた場合、その内容を一致させるためのコピー動作を説明するためのシリングマップを示す図である。

第1図において、OS101はチャネル111、112を介して、OS102はチャネル113、114を介してディスク制御装置121、122

スク装置131、132、141、142内に夫々設けてよい。

以下、OS101がディスク装置131に対して入出力動作を実行する場合について説明する。

OS101は、チャネル111、ディスク制御装置121を介してメモリ151をアクセスするが、この場合、第3図に示すコマンド・チエインを使用してメモリ151をアクセスする。本発明のために新たなコマンドとして、CHECK TABLE 301及びREAD TABLE 305が用意されている。CHECK TABLE 301は、メモリ151または152中の管理状態情報がOSの期待する条件を満足するか否か、すなわち、入出力装置に対応して自OS内に有する管理状態情報とメモリ151または152内の管理状態情報とが一致するか否かをテストするコマンドである。READ TABLE 305は、メモリ151または152中の管理状態情報を読み出すコマンドである。

OS101がディスク装置131に対して入出力動作を実行するために、OS101は、まず前

と接続され、ディスク制御装置121及び122は、夫々複数のディスク装置131、132および141、142に接続されている。また、ディスク制御装置121および122は、自己が制御するディスク装置の各々の管理状態を管理状態情報として記憶するメモリ151および152を備えている。

この管理状態情報は、OSが各ディスク装置の管理を行うために必要とする情報であり、通常、OS内に記憶されているが、この情報のコピーをメモリ151、152内にも存在させることになる。メモリ151にはディスク制御装置121配下の全ディスク装置131、132に対する管理状態情報が、メモリ152にはディスク制御装置122配下の全ディスク装置141、142に対する管理状態情報が夫々格納記憶されている。このメモリ151、152は、高速アクセス可能なものがよく、一般的には、半導体メモリが適している。また、このメモリ151、152は、ディスク制御装置121、122内ではなく、各ディ

述のコマンド・チエインのCHECK TABLE 301を実行し、自己の有するディスク装置131に対応する管理状態情報と、メモリ151内のディスク装置131に対応する管理状態情報とが一致しているか否かのチェックを実行する。一致が得られた場合、OSのコマンド実行は、公知のジャンプ先変更コマンドTIC303により、ユーザ・コマンド・チエイン304にジャンプされ、このユーザ・コマンド・チエイン304に対して行われることになる。このユーザ・コマンド・チエイン304は、通常用いられている入出力動作を実行するコマンド・チエインであり、OS101は、このコマンドを実行することにより、ディスク装置131に対する所定の入出力動作を完了する。

OS101が、コマンドCHECK TABLE 301を実行し、メモリ151内及び自OS内のディスク装置対応の管理状態情報の不一致を検知した場合、OS101のコマンドの実行は、ジャンプ先変更コマンドTIC302により、コマンドREAD TABLE 305にジャンプされ、このコマンド305

特開昭62-274448 (4)

に対して行われる。OS101は、このコマンド305の実行により、ディスク装置131の管理状態情報をメモリ151より読み出し、読み出された管理状態情報を解析することにより、その後の適切な処理を実行することができる。また、メモリ151から読み出したディスク装置131の管理状態情報は、OS101内のディスク装置131の管理状態情報としてOS101内に取り込まれる。

以上の動作は、OS102がディスク装置131に対する入出力動作を実行する場合も、またOS101, 102がディスク装置131以外のディスク装置に対する入出力動作を実行する場合も全く同様に行われる。前述した入出力装置としてのディスク装置の制御において、管理状態情報の変更は、OSからのソフトウェアによる指令により行われてもよく、或いは、通常の入出力動作の一部として行われてもよい。また、管理状態情報の変更の連絡の契機は、入出力装置自身がとつてもよいし、別のソフトウェアによる指令によるものであつてもよい。

このように、OS101, 102は、コマンド

を保持するものである。

いま、第1図において、ディスク装置131と142とがお互いに二重化されているものとする。

この場合、OS101または102からの書き込み動作は、OS101または102がディスク装置131, 142の両方に入出力命令を発して実行され、読み出し動作は、OS101または102がディスク装置131, 142の何れか一方に入出力命令を発して実行されることになる。このような二重化されたディスク装置131, 142を複数のOS101, 102から使用する場合には、次のような問題が生じる。すなわち、OS101がディスク装置131にデータを書き込んだ後、ディスク装置142に同一のデータを書き込む処理を実行しようとしたとき、ディスク装置142に障害が発生すると、以後OS101は、ディスク装置131のみを使用し、ディスク装置142を論理的に閉塞し使用しなくなり、第1図に示す情報処理システムは、事实上片肺運転状態となる。この状態で、OS102が、二重化

C H E C K T A B L E 3 0 1 を実行することにより、メモリ151, 152内の各ディスク装置の管理状態情報が期待する条件を満足する場合のみ、各ディスクに対する所定の入出力動作を実行し、満足しない場合、管理状態情報の変更を検知して適切な処理を実行できる。

以上により、OS101, 102は、共通に使用するディスク装置の管理状態の変更を、特別にOS101, 102相互間で連絡を取り合うことなく検知することができる。そして、すでに述べたように、メモリ151, 152には、高速アクセス可能な半導体メモリ等を用いており、また、OSは、通常の入出力動作の一部としてディスク装置の管理状態の変更を検知できるので、システム・スループットの低下を抑止できる。

次に、二重化されたディスク装置に対して本発明を適用した場合について説明する。

一般に、入出力装置の二重化は、情報処理システム全体の信頼性向上のために行われる手段であり、二重化された入出力装置は、必ず同一の情報

されたディスク装置の一方のディスク装置142からデータの読み出し動作を行うと、OS102は、未だディスク装置の前述の状態を検知しておらず、ディスク装置142からOS101が書き込んでいない古いデータを読み出して使用するという問題が発生する。

本発明によれば、このような問題も解消することができる。以下に、第2図及び第4図も参照して、この場合の動作を説明する。第2図に示す1個のディスク装置に対する管理状態情報のメモリマップにおいて、二重化状態フラグ201は、そのディスク装置が二重化されているか否かを、閉塞状態フラグ202は、そのディスク装置が論理的に閉塞されているか否かを、コピー状態フラグ203は、そのディスク装置が片肺運転から両肺運転に移行するためのコピー動作の実行中であるか否かを夫々示しており、コピー中シリンド番号204は、コピー動作が現在行われているシリンド番号を示している。また、入出力動作保留中フラグ205は、そのディスク装置がOSからの

特開昭62-274448 (5)

入出力動作を一旦保留したことを示している。ディスク制御装置内に設けられたメモリ151,152には、必要に応じて上記以外の管理状態情報を記憶格納することが可能である。

いま、OS101がディスク装置131にデータを書き込んだ後、ディスク装置142に同一のデータを書き込む動作の実行を行うものとする。すでに述べたように、OS101は、第3図に示すコマンド・チエインを実行し、ディスク制御装置121内のメモリ151のディスク装置131に対応する管理状態情報を読み出し、該情報が期待する条件を満足する場合にディスク装置131にデータを書き込むことができる。さらに、OS101は、同様にしてディスク制御装置122内のメモリ152のディスク装置142に対応する管理状態情報を読み出して、ディスク装置142にディスク装置131に書き込んだデータと同じデータを書き込もうとする。このとき、ディスク装置142に障害が発生すると、OS101は、この書き込み動作を中止し、メモリ152のディ

スク装置142に対応する管理状態情報の閉塞状態フラグ202を「オフ」から「オン」に変更する。これにより、第1図の情報処理システムは、両肺運転状態から片肺運転状態に変更され、二重化されたディスク装置の一方のディスク装置142が閉塞状態とされる。

その後、OS102がディスク装置142のデータを読み出すため、第3図のコマンド・チエインを実行すると、OS102は、ディスク装置142の閉塞状態フラグ202が「オフ」であることを期待しているため、CHECK-TABLE301で条件不満足となり、ユーザ・コマンド・チエイン304は実行されないことになる。これにより、OS102は、OS101が書き込んでいない古いデータを使用しないで済み、また、OS101によつてディスク装置142が閉塞させられ、管理状態情報が変更されたことを検知でき、通常の入出力動作の一部として、変更された管理状態情報を自OS内に取り込み自己が有する管理状態情報を更新することができる。

第1図に示す情報処理システムが、このような片肺運転を続けると、OS101及び102は、ディスク装置131のみを使用することになり、ディスク装置131と142の記録内容は異なるものとなる。このため、片肺運転状態を両肺運転状態に戻すためには、ディスク装置131の内容をディスク装置142にコピーする必要がある。本発明においては、メモリ151,152を使用することにより、効率のよいコピー動作を実行することができ、かつ、コピー動作中のディスク装置の閉塞を効果的に行うことができる。

第4図は、コピー動作を説明するためのディスク装置131,142のシリンドラマップを示しており、ディスク装置131,142は、夫々領域401,402で示されるル本のシリンドラ(シリンドラ番号0～ル-1)を有する。コピー動作は、シリンドラ番号0から始まつてシリンドラ番号の昇順に実行されるものとし、現在シリンドラ番号エの領域413から領域414へのコピーを実行しているものとすると、シリンドラ番号0～エ-1の領域

411,412は、コピー済みであり、ディスク装置131及び142のこの領域の内容は、同一となる。シリンドラ番号エ+1～ル-1の領域415及び416は末コピー領域であり、その内容は異なる可能性がある。このコピー動作は、OS101または102によつて実行されるが、ディスク装置142を閉塞したままコピーを実行すると、このコピー動作中に、コピー動作以外の通常の入出力動作によつて、ディスク装置131のコピー済みの領域411の内容が書き換えられる場合が生じる。この場合、領域411と412の内容が異なる状態になり、コピー動作が終了しても両肺運転状態に移行できることになる。これを防ぐために、コピー中は、コピー動作以外の通常の入出力動作による書き込み動作を禁止すればよいが、そうすると、システム・スループットの低下を招いてしまう。

本発明によれば、コピー動作中は、ディスク装置131,142の全体を閉塞するのではなく、各ディスク装置の現在コピー中の領域413,

特開昭62-274448 (6)

414及びディスク装置142の末コピーの領域416のみが閉塞される。これにより、両ディスク装置のコピー済領域411, 412に対するコピー動作中のコピー動作以外の通常の入出力動作は、両肺運転によりその実行が可能とされ、データ内容の同一性が維持される。コピー中の領域413, 414に対するコピー動作以外の通常の入出力動作は、そのシリンドのコピーが終了する迄その実行が保留される、未コピー領域415に対するコピー動作以外の通常の入出力動作は、ディスク装置131のみの片肺運転によりその実行が可能とされる。

OS101がディスク装置131からディスク装置142へのコピー動作を実行中、OS102がコピー動作以外の通常の入出力動作を実行しようとする場合のために、OS101は、ディスク制御装置121, 122内のメモリ151, 152を用いてディスク装置131, 142の閉塞と閉塞解除を次の様に実行する。OS101は、コピー動作に先立つてメモリ151中のディスク装置

OS102に対して入出力刺込みにより、シリンド番号「0」に対する閉塞が解除されたことを報告し、メモリ151中の入出力動作保留中フラグ205を「オフ」にする。OS101は、このようにして順次コピー中のシリンドの閉塞とコピー終了シリンドの閉塞解除を繰返すことにより、ディスク装置131の全領域401の内容をディスク装置142の全領域402へコピーする。全てのコピーが終了すると、OS101は、メモリ151中のディスク装置131に対応する管理状態情報のうちコピー状態フラグ203を「オフ」にし、メモリ152中のディスク装置142に対応する管理状態情報のうちコピー状態フラグ203と閉塞状態フラグ202を「オフ」にする。これにより第1図に示した情報処理システムは、両肺運転状態にされる。

OS101がシリンド番号「x」の領域413を領域414にコピー動作を実行中に、OS102がディスク装置131, 142に対してコピー動作以外の通常の入出力動作を実行しようとする場

131に対応する管理状態情報およびメモリ152中のディスク装置142に対応する管理状態情報のうち、コピー状態フラグ203を「オン」とし、コピー中シリンド番号204のシリンド番号を「0」とする。OS101は、シリンド番号「0」のデータをディスク装置131から142へコピーし、コピー終了後コピー中シリンド番号204のシリンド番号を1だけ増加させ「1」とし、このシリンドに対するコピー動作を継続するとともに、シリンド番号「0」に対する閉塞の解除をディスク制御装置121または122に指示する。OS101がシリンド番号「0」のデータをコピー中に、OS102がシリンド番号「0」に対して入出力動作を起すと、ディスク制御装置121内のメモリ151中の入出力動作保留中フラグ205は「オン」とされ、OS102からのこの入出力動作は、保留される。ディスク制御装置121は、OS101からシリンド番号「0」に対する閉塞の解除が指示されると、メモリ151中の入出力動作保留中フラグ205を調べ、これが「オン」であれば、

合、OS102は、第3図のコマンド・チエインを使用すればよい。OS102は、ディスク装置131に対しては、コピー状態フラグ203が「オン」でコピー中シリンド番号204が、OS102が実行しようとしている入出力動作のシリンド番号と異なることを期待して、またディスク装置142に対しては、コピー状態フラグ203が「オン」でコピー中シリンド番号204が、OS102が実行しようとしている入出力動作のシリンド番号より大きいことを期待してCHECK TABLE301を実行する。これにより、ディスク装置131のコピー中領域413に対する入出力動作は実行されないが、コピー済領域411と未コピー領域415に対する入出力動作は実行され、また、ディスク装置142のコピー中領域414と未コピー領域416に対する入出力動作は実行されず、コピー済領域412に対する入出力動作は実行されることになる。

前述した。OS101によるディスク装置131からディスク装置142へのコピー動作の実行中、

特開昭62-274448 (7)

OS 102 がディスク装置 131 または 142 に対してコピー動作以外の通常の入出力動作実行しようとして、その入出力動作が実行されず保留されると、ディスク制御装置 121 または 122 は、メモリ 151 または 152 内の入出力動作保留フラグ 205 を「オン」とする。コピー動作の進行によりコピー中のシリンド番号が変化することにより、ディスク制御装置 121 または 122 は、メモリ 151 または 152 内の入出力動作保留中フラグ 205 が「オン」になつている場合に、OS 102 に対してディスク装置 131 または 142 の閉塞の解除を入出力割込により報告することは既に述べたとおりである。

以上説明した実施例では、ディスク装置 131 と 142 が二重化されているとして説明したが、ディスク装置 132 と 141 とが二重化されている場合も同様である。また、二重化されたディスク装置間のコピー動作においてそのコピー範囲をディスク装置全体としたが、ディスク装置の領域の一部に限つてコピー動作を実行することも可能

度の閉塞を効果的に実現できるという効果がある。

4. 図面の簡単な説明

第1図は本発明の一実施例を示すブロック図、第2図はメモリマップの一例を示す図、第3図はコマンド・チエインを示す図、第4図はシリンドマップを示す図である。

101, 102 ……オペレーティング・システム、111～114 ……チャネル、121, 122 ……ディスク制御装置、131, 132, 141, 142 ……ディスク装置、151, 152 ……メモリ。

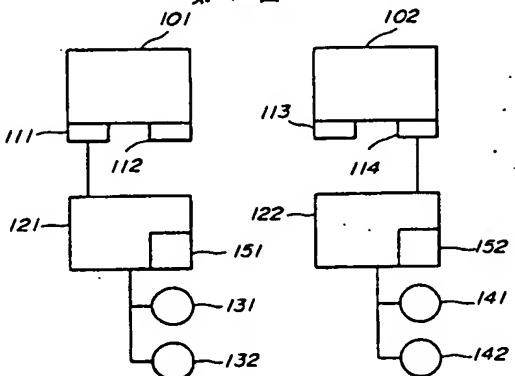
代理人弁理士 武 路次郎(外1名)

である。さらに、コピー中であるとして閉塞するディスク装置の領域をシリンド単位としたが、ディスク制御装置内に設けるメモリの容量を増加することにより、ディスク装置の磁気ヘッド単位、すなわちトラック単位としてもよい。

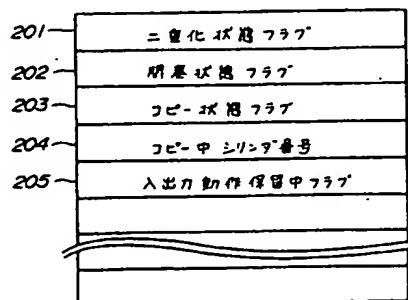
〔発明の効果〕

本発明によれば、入出力装置を複数のOSから共用する場合に、OS間に例えばチャネル間結合装置のような特別の連絡手段を設けて、OS相互間で特別に連絡を取り合うことなく、入出力装置の管理状態の変更を可能にし、通常の入出力動作の一部として、他のOSによる入出力装置の管理状態の変更を認識できる。このため、OS間の連絡に要するオーバヘッドを大幅に削減でき、システム・スループットの低下を招くことなく、また運用上の制限を受けることなく、ファイルの共用が可能になるという大きな効果がある。さらに本発明によれば、入出力装置が二重化されている場合のコピー動作においても、システム・スループットの低下を招く事なくコピー動作時の入出力

第1図

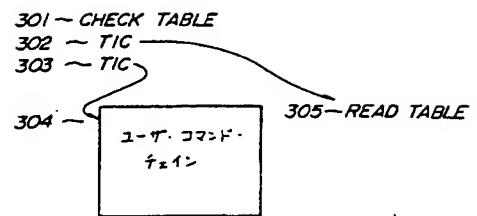


第2図



特開昭62-274448 (8)

第3圖



第 4 図

